和歌山大学経済学会『研究年報』第14号（2010年）pp.533−542　ISSN 1880−490X

適応的システムと確率的環境の相互作用

The Interaction between Adaptive System and Stochastic Environment

竹　内　昭　浩
Akihiro TAKEUCHI

1．はじめに

本稿においては，確率的環境内における適応的システムの動作について議論するが，これは，未知の環境内で適応的に行動するためにはどのような行動基準が必要であり，また如何なる能力を必要とするのかを明らかにしようとする試みであり，これまでにも多くの研究者により考察されている。(1)

さて，次の様な状況を考えてみよう。

二台のスロットマシーンが並んでおり，この機械を相手に賭けをするとして何れの機械と勝負すればよいのだろうか。当然当たりの出る割合の多いほうと勝負をしたいのだが，どの様にすればそれがわかるのであろうか。しかし，これはやってみなければわからない。そこでどういう作戦でもって早く正確に当たりの出る割合の多い機械を発見するかという問題が生じる。これは二つ手の悪漢問題とも呼ばれている学習問題である。

ところで，このような状況はスロットマシーンの場合だけではなく，不確かな環境下で色々と試行錯誤しながら行動し，その環境に適応してゆく問題に共通する性質を有している。そこで，先の二つ手の悪漢問題を少し抽象化して次のように考えてみよう。二つの機械のうち，第一のそれで勝つ確率をq_1，負ける確率をp_1（$q_1+p_1=1$）とし，第二のそれで勝つ確率をq_2，負ける確率をp_2（$q_2+p_2=1$）とする。この確率を知らずに，色々と試しながらどちらの機械と勝負するかを決めてゆくのである。このような確率的環境はランダム環境と呼ばれ，以下のように定義される。

［定義１］ランダム環境$C=C(p_1, p_2)$とはその環境内で行動するシステムの出力o_i（$i=1$ or 2）に対して確率p_iでもって損失（システムの負け）を，確率q_i（$p_i+q_i=1$）でもって利得（システムの勝ち）を呈する環境である。

（1）Tsetlin［T- 3］, Fu［F- 1］, 竹内［T- 1］, 竹内［T- 2］等

次節においてCover-Hellmannら[(2)]によるシステムの上記ランダム環境内での動作にについて議論したのち，3節では彼らのシステムを結合したシステムTを提案し，これらシステムの環境下での振る舞いを比較検討する。

2．適応的システムΣ_1，Σ_2

次のようなシステムΣ_1を考えよう。

1．状態数はすべてでn＋2個とし，それぞれr_0，r_1，r_2，……r_{n-1}，r_n，r_{n+1}とする。
2．状態r_0では，確率（1－b）で戦略o_1を，確率bで戦略o_2を取る。
3．状態r_{n+1}では，確率（1－b）で戦略o_2を，確率bで戦略o_1取る。
4．その他の状態では確率½で戦略o_1またはo_2を取る。
5．状態r_1からr_{n+1}までの状態r_jで，o_1を選んで勝った時には状態r_{j-1}に推移する。
6．状態r_0からr_nまでの状態r_jで，o_2を選んで勝った時には状態r_{j+1}に推移する。
7．その他の場合には状態推移はない。

このシステムと先に定義したランダム環境Cとの相互作用を解析してみよう。この相互作用系が有限マルコフ過程となることは明らかである。そこでこのマルコフ過程の状態推移行列をPとおくと，

$$P=\begin{pmatrix} (1-bq_2) & bq_2 & & & & \\ \frac{q_1}{2} & \frac{p_1+p_2}{2} & \frac{q_2}{2} & & O & \\ & \ddots & \ddots & \ddots & & \\ & & \frac{q_1}{2} & \frac{p_1+p_2}{2} & \frac{q_2}{2} & \\ & & & \ddots & \ddots & \ddots \\ & O & & \frac{q_1}{2} & \frac{p_1+p_2}{2} & \frac{q_2}{2} \\ & & & & bq_1 & (1-bq_1) \end{pmatrix} \tag{1}$$

となる。またこの系はエルゴード的であるので，定常状態時におけるシステムΣ_1の状態がr_jである確率をπ_jとすると，

$$\pi_0=(1-bq_2)\pi_0+\frac{q_1}{2}\pi_1$$

$$\pi_1=bq_2\pi_0+\frac{p_1+p_2}{2}\pi_1+\frac{q_1}{2}\pi_2$$

$$\vdots$$

（2）このことについては，甘利[A-1]の解説等参照

$$\pi_j = \frac{q_2}{2}\pi_{j-1} + \frac{p_1+p_2}{2}\pi_j + \frac{q_1}{2}\pi_{j+1} \tag{2}$$

$$\vdots$$

$$\pi_n = \frac{q_2}{2}\pi_{n-1} + \frac{p_1+p_2}{2}\pi_n + bq_1\pi_{n+1}$$

$$\pi_{n+1} = (1-bq_1)\pi_{n+1} + \frac{q_2}{2}\pi_n$$

となる。ここで

$$\pi_j = x_1^j \tag{3}$$

と置くと

$$2x_1^j = q_2 x_1^{j-1} + (p_1+p_2)x_1^j + q_1 x_1^{j+1} \tag{4}$$

となり，これより

$$q_1 x_1^2 - (q_1+q_2)x_1 + q_2 = 0 \tag{5}$$

$$(q_1x_1 - q_2)(x_1 - 1) = 0 \tag{6}$$

となる。ゆえに，K_1，L_1を任意定数として，一般解π_jは

$$\pi_j = K_1\left(\frac{q_2}{q_1}\right)^j + L_1 \tag{7}$$

で与えられる。また

$$2bq_2\pi_0 = q_1\pi_1 \tag{8}$$

より

$$\pi_0 = \frac{q_1}{2bq_2}\pi_1 \tag{9}$$

となり，

$$\pi_1 = \frac{q_1}{2}\pi_1 + \frac{p_1+p_2}{2}\pi_1 + \frac{q_1}{2}\pi_2 \tag{10}$$

を得，これを整理すると

$$q_1\pi_2 - q_2\pi_1 = 0 \tag{11}$$

となる。これにさきほどのπ_jを代入すると

$$q_1(K_1\frac{q_2^2}{q_1^2}+L_1)-q_2(K_1\frac{q_2}{q_1}+L_1)=0 \tag{12}$$

となり，これより

$$L_1=0 \tag{13}$$

となる。ここで

$$\alpha_1=\frac{q_2}{q_1} \tag{14}$$

と置くと，

$$\sum_{j=0}^{n+1}\pi_j=1 \tag{15}$$

より

$$K_1\left\{\frac{1}{2b}(1+\alpha_1^{n+1})+\frac{\alpha_1-\alpha_1^{n+1}}{1-\alpha_1}\right\}=1 \tag{16}$$

となり，これより K_1 は

$$K_1=\frac{2b}{(1+\alpha_1^{n+1})+2b(\frac{\alpha_1-\alpha_1^{n+1}}{1-\alpha_1})} \tag{17}$$

となる。状態 r_0 では確率（１－b）で戦略 o_1 を，確率bで o_2 を取り，状態 r_{n+1} では確率（１－b）で戦略 o_2 を，確率bで o_1 を取り，その他の状態では確率½で戦略 o_1 または o_2 を取るので，この相互作用系の定常状態時におけるシステム Σ_1 の損失の期待値を $\sigma(\Sigma_1;C)$ と置くと

$$\sigma(\Sigma_1;C)=\frac{K_1}{2b}\{(1-b)p_1+bp_2\}+K_1\frac{p_1+p_2}{2}\frac{\alpha_1-\alpha_1^{n+1}}{1-\alpha_1}+\frac{K_1}{2b}\alpha_1^{n+1}\{(1-b)p_2+bp_1\} \tag{18}$$

を得る。この $\sigma(\Sigma_1;C)$ は明らかにbに対しては単調増加，nに対しては単調減少となっている。そこで，bを十分小さくし，nを十分大きくした時その値は

$$\min\{C(p_1, p_2)\}$$

に限りなく近づいていくこと，つまりその条件下でシステム Σ_1 がランダム環境に対して適応的であることが分かる。

ところで，上記のシステム Σ_1 は利得を得たときにのみ状態の推移が起こっているが，逆に損失を被ったときにのみ状態推移を行うシステムも考えることができる。このシステムを Σ_2 と名付けよう。Σ_2 は出力に関しては Σ_1 とまったく同じであるが，状態推移に関して以下のようになる。

1．状態r_0では，確率bで戦略o_1を取り負けたとき状態o_2へ推移する。
2．状態r_{n+1}では，確率bで戦略o_2を取り負けたとき状態o_nへ推移する。
3．状態r_1からr_nまでの状態r_jでo_1を取り負けたときには状態r_{j+1}に推移する。
4．状態r_1からr_nまでの状態r_jでo_2を取り負けたときには状態r_{j-1}に推移する
5．その他の場合には状態推移はない。

このシステムΣ_2が，環境C内において十分長い時間相互作用したのちの損失の期待値$\sigma(\Sigma_2;C)$に関しては，システムΣ_1の場合とほぼ同様にして解くことができ，

$$\sigma(\Sigma_2;C)=\frac{K_2}{2b}\{(1-b)p_1+bp_2\}+K_2\frac{p_1+p_2}{2}\frac{\alpha_2-\alpha_2^{n+1}}{1-\alpha_2}+\frac{K_2}{2b}\alpha_2^{n+1}\{(1-b)p_2+bp_1\} \quad (19)$$

ただし，

$$\alpha_2=\frac{p_1}{p_2} \quad (20)$$

$$K_2=\frac{2b}{(1+\alpha_2^{n+1})+2b\left(\frac{\alpha_2-\alpha_2^{n+1}}{1-\alpha_2}\right)} \quad (21)$$

となる。この$\sigma(\Sigma_2;C)$も明らかに，bに対しては単調増加，nに対しては単調減少であり，bを十分小さくし，nを十分大きくした時のその値は，

$$\min\{C(p_1,\ p_2)\}$$

に限りなく近づいていくこと，つまりシステムΣ_2もランダム環境に対して適応的であることが分かる。

3．適応的システムT

本節では，2．において考察した二つのシステムΣ_1およびΣ_2を合体したシステムTを提案し，このシステムとΣ_1およびΣ_2との動作の比較を行う。

次のようなシステムTを考えよう。

1　状態数はすべてでn＋2個とし，それぞれr_0，r_1，r_2，……r_{n-1}，r_n，r_{n+1}とする。

2　状態r_0では

　2．1　確率bで戦略o_1を取り，その結果損失を受けた時は状態r_1へ推移し，利得を受けた時は状態推移しない。

2．2　確率bで戦略o_2を取り，その結果利得を受けた時は状態r_1へ推移し，損失を受けた時は状態推移しない。

2．3　残りの確率で戦略o_1を取り，状態推移はしない。

3　状態r_{n+1}では

3．1　確率bで戦略o_2を取り，その結果損失を受けた時は状態r_nへ推移し，利得を受けた時は状態推移しない。

3．2　確率bで戦略o_1を取り，その結果利得を受けた時は状態r_nへ推移し，損失を受けた時は状態推移しない。

3．3　残りの確率で戦略o_2を取り，状態推移はしない。

4　その他の状態r_jでは確率½で戦略o_1またはo_2を取り，

4．1　戦略o_1を選んで利得を得たとき，およびo_2を選んで損失を受けたときは，状態r_{j-1}に推移する。

4．2　戦略o_2を選んで利得を得たとき，およびo_1を選んで損失を受けたときは，状態r_{j+1}に推移する。

このシステムと先に定義したランダム環境Cとの相互作用を解析してみよう。この相互作用系が有限マルコフ過程となることは明らかである。そこでこのマルコフ過程の状態推移行列をPとおくと，

$$P=\begin{pmatrix}
(1-b(q_2+p_1)) & b(q_2+p_1) & & & & & \\
\frac{q_1+p_2}{2} & 0 & \frac{q_2+p_1}{2} & & & O & \\
 & \ddots & \ddots & \ddots & & & \\
 & & \frac{q_1+p_2}{2} & 0 & \frac{q_2+p_1}{2} & & \\
 & & & \ddots & \ddots & \ddots & \\
 & O & & & \frac{q_1+p_2}{2} & 0 & \frac{q_2+p_1}{2} \\
 & & & & & b(q_1+p_2) & (1-b(q_1+p_2))
\end{pmatrix} \tag{22}$$

となる。またこの系もエルゴード的であるので，定常状態時におけるTの状態がr_jである確率をπ_jとすると，

$$\pi_o=(1-b(q_2+p_1))\pi_o+\frac{q_1+p_2}{2}\pi_1$$

$$\pi_1 = b(q_2+p_1)\pi_o + \frac{q_1+p_2}{2}\pi_2$$

$$\vdots$$

$$\pi_j = \frac{q_2+p_1}{2}\pi_{j-1} + \frac{q_1+p_2}{2}\pi_{j+1} \tag{23}$$

$$\vdots$$

$$\pi_n = \frac{q_2+p_1}{2}\pi_{n-1} + b(q_1+p_2)\pi_{n+1}$$

$$\pi_{n+1} = (1-b(q_1+p_2))\pi_{n+1} + \frac{q_2+p_1}{2}\pi_n$$

となる。先ほどのΣ_1の解析と同様に

$$\pi_j = x^j \tag{24}$$

と置くと，

$$((q_1+p_2)x-(q_2+p_1))(x-1)=0 \tag{25}$$

となる。ゆえに，K，Lを任意定数として，一般解π_jは

$$\pi_j = K\left(\frac{q_2+p_1}{q_1+p_2}\right)^j + L \tag{26}$$

で与えられる。また

$$2b(q_2+p_1)\pi_0 = (q_1+p_2)\pi_1 \tag{27}$$

より

$$\pi_0 = \frac{(q_1+p_2)}{2b(q_2+p_1)}\pi_1 \tag{28}$$

となり，

$$\pi_1 = \frac{(q_1+p_2)}{2}\pi_1 + \frac{(q_1+p_2)}{2}\pi_2 \tag{29}$$

を得，これを整理すると

$$(q_2+p_1)\pi_2 - (q_1+p_2)\pi_1 = 0 \tag{30}$$

となる。これにさきほどのπ_jを代入すると

$$(q_1+p_2)\left(K\frac{(q_2+p_1)^2}{(q_1+p_2)^2}+L\right) - (q_2+p_1)\left(K\frac{(q_2+p_1)}{(q_1+p_2)}+L\right) = 0 \tag{31}$$

となり，これより

$$L=0 \tag{32}$$

を得る。ここで

$$\alpha=\frac{q_2+p_1}{q_1+p_2} \tag{33}$$

と置くと

$$\sum_{j=0}^{n+1}\pi_j=1 \tag{34}$$

より

$$K\left\{\frac{1}{2b}(1+\alpha_1^{n+1})+\frac{\alpha-\alpha_1^{n+1}}{1-\alpha}\right\}=1 \tag{35}$$

となり，これより K は

$$K=\frac{2b}{(1+\alpha^{n+1})+2b\left(\dfrac{\alpha-\alpha^{n+1}}{1-\alpha}\right)} \tag{36}$$

となる。状態 r_0 では確率（1－b）で戦略 o_1 を，確率 b で o_2 を取り，状態 r_{n+1} では確率（1－b）で戦略 o_2 を，確率 b で o_1 を取り，その他の状態では確率½で戦略 o_1 または o_2 を取るので，この相互作用系の定常状態時におけるシステム T の損失の期待値を τ(T；C）と置くと

$$\tau(T;C)=\frac{K}{2b}\{(1-b)p_1+bp_2\}+K\frac{p_1+p_2}{2}\frac{\alpha-\alpha^{n+1}}{1-\alpha}+\frac{K}{2b}\alpha_1^{n+1}\{(1-b)p_2+bp_1\} \tag{37}$$

を得る。この τ(T；C）は明らかに b に対しては単調増加，n に対しては単調減少となっている。そこで，b を十分小さくし，n を十分大きくした時その値は

min $\{C(p_1,\ p_2)\}$

に限りなく近づいていくこと，つまりその条件下でシステム T がランダム環境に対して適応的であることが分かる。

ここで，システム T とシステム Σ_1，Σ_2 の損失の期待値を比較してみよう。図 1，図 2 はそれぞれ C＝C（0.6，0.8）および C＝C（0.2，0.4）という二つのランダム環境における損失の期待値を b＝0.01 として状態数の変化に対してプロットしたものであり，図 3，図 4 はそれぞれ同じ環境下で，n＝5 として b の変化に対してプロットしたものである。

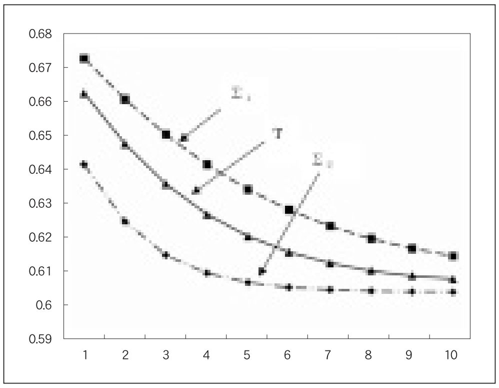

図 1

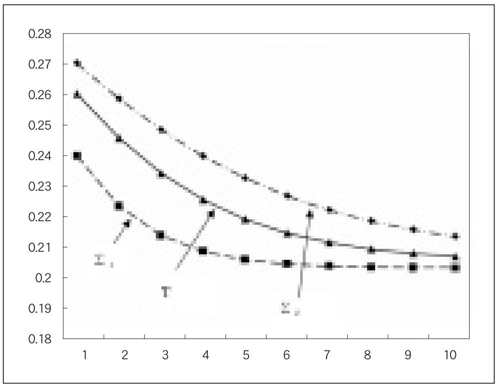

図 2

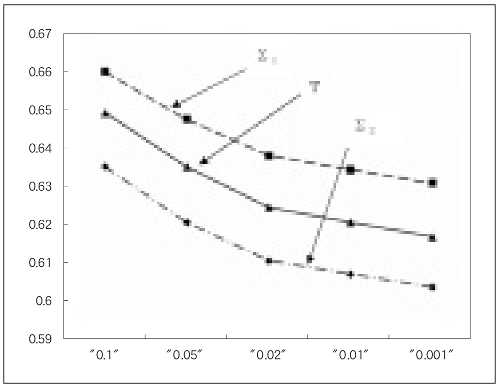

図 3

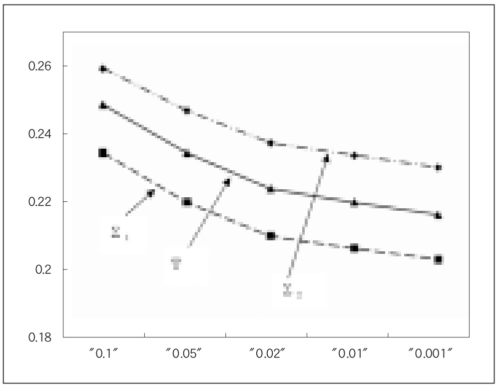

図 4

また，$p_1+p_2=1$のとき，

$$\alpha_1=\alpha_2=\alpha=\frac{p_2}{p_1}$$

となり，

$$\sigma(\Sigma_1;C)=\sigma(\Sigma_2;C)=\tau(T;C)$$

が成り立つ。このことと式（18），（19），（37）さらに図1，2，3，4の比較から$p_1+p_2>1$のときには

$$\sigma(\Sigma_1;C)<\tau(T;C)<\sigma(\Sigma_2;C)$$

となり，$p_1+p_2<1$のときには

$$\sigma(\Sigma_1;C)>\tau(T;C)>\sigma(\Sigma_2;C)$$

となり，いずれの場合においてもシステムTはΣ_1，Σ_2の中間的な適応性を示しており，より広い範囲での平均的な適応能力を持っていることが分かる。

4．まとめ

本稿において，Cover-Hellmanによるシステムについて考察した後，それらのシステムを結合したシステムTを提案した。そして，このシステムが先のシステムの中間的な適応能力を持ち，より広い範囲での適応能力を有することを示した。

ところで，これらシステムが環境内でどのような速度で適応していくかという適応速度の問題や，このような適応的システムが複数個集まって行動する姿の観察など，まだまだ興味深い問題が残されているが，それらについてはまた稿を改めて議論したい。

［A- 1］甘利俊一：「機械系と生体系にみる，学習と自己組織系の理論」，数理科学，195，pp.18-22，（1979）

［F- 1］Fu, K. S., Li, T. J.：“Formulation of learning automata and automata games”, Information Sciences, 1, pp.237-256（1969）

［T- 1］竹内昭浩：人間行動とオートマトン　白桃書房（1984）

［T- 2］竹内昭浩：「適応的システムの確立環境内で動作とその順応速度」経済理論，334，pp.145-160（2006）

［T- 3］Tsetlin, M. L.：“On the behavior of finite automata in random media”, Automation and Remote Control, 22, pp.1345-1354（1961）